

FAIR Biomedical Research Software (FAIR-BioRS) guidelines v1.0.0

Step 1: Follow standards and best practices during development of the software

Step 1.a. Follow general standards and best practices for scientific software, such as working from a version control system (GitHub, Bitbucket, GitLab), having code level documentation (in code comments, description in the headers), and recording dependencies in a requirement.txt file or a README file. It is recommended to read the “Good enough practices in scientific computing”¹ and the “General guidelines for biomedical software development”².

Step 1.b. Follow language-specific standards and best practices. They will depend on the development stack used (e.g., the Python Developer’s Guide³ for Python code or Google’s R Style Guide⁴ for R code).

Step 1.c. Follow standards and best practices for documenting your software. It is suggested to maintain the documentation in a README.md or README.txt file located in the root directory of the software. The following aspects must be documented: inputs and outputs of the software, parameters and data required to run the software, the standards applied, and how to contribute to the software. It is suggested to read the paper on the “Ten simple rules for documenting scientific software”⁵.

Step 2: Collect files to share

Include the source code when possible and also include executables if applicable.

Step 3: Include metadata files

Include both a codemeta.json and CITATION.cff metadata files in the root directory of the software. The CodeMeta generator^{6,7} and CFF file initializer^{8,9} available to help prepare both.

Provide, at least, the following fields in the codemeta.json file: Software name (“name”), Software description/abstract (“description”), Authors (“givenName”, “familyName”) with their Organization name (“affiliation”), Keywords (“keywords”), Programming Language (“programmingLanguage”), Release date (“dateModified”), License used (“license”)

Provide, at least, the following fields in the CITATION.cff file: Authors (“given-names”, “family-names”) with their Organization name (“affiliation”), Software description/abstract (“abstract”), Keywords (“keywords”), License (“license”), Release date (“date-released”)

Step 4: Choose a license

Include the license terms in a LICENSE file located in the root directory of the software. While the FAIR4RS principles do not require research software to be open-source, it is highly recommended using a license approved by the Open Source Initiative (OSI)¹⁰. Amongst those licenses, it is encouraged to use the permissive MIT¹¹ or Apache 2.0¹² licenses.

Step 5: Share software on a repository

Step 5.a. Share on a language-specific (e.g., PyPI¹³ or Conda¹⁴ for Python packages, CRAN for R packages¹⁵, Dockstore for Docker-based tools¹⁶) or biomedical-specific (e.g., ModelDB for computational neuroscience models^{17,18} and Bioconductors for R-packages aimed at the analysis of genomics data^{19–21}) repository if applicable. The Registry of Research Data Repositories can be used to find a suitable repository²².

Step 5.b. Share on Zenodo or Figshare.

Step 6: Register software on a registry

Register the software on bio.tools or on the Research Resource Identifiers (RRID) Portal.

References

1. Wilson, G. *et al.* Good enough practices in scientific computing. *PLoS Comput. Biol.* 13, e1005510 (2017).
2. Silva, L. B., Jimenez, R. C., Blomberg, N. & Oliveira, J. L. General guidelines for biomedical software development. *F1000Research* vol. 6 273 (2017).
3. Python Developer's Guide. *Python.org* <https://www.python.org/dev/>.
4. Google's R style guide. *styleguide* <https://google.github.io/styleguide/Rguide.html>.
5. Lee, B. D. Ten simple rules for documenting scientific software. *PLoS Comput. Biol.* 14, e1006561 (2018).
6. CodeMeta generator. <https://codemeta.github.io/codemeta-generator/>.
7. *codemeta-generator: This repository contains a (client-side) web application to generate Codemeta documents (aka. codemeta.json).* (Github).
8. cffinit: a web form to initialize CITATION.cff files.
<https://citation-file-format.github.io/cff-initializer-javascript/#/>.
9. *cff-initializer-javascript: Web form to initialize CITATION.cff files.* (Github).
10. Licenses & standards. <https://opensource.org/licenses>.
11. The MIT license. <https://opensource.org/licenses/MIT>.
12. Apache license, version 2.0. <https://opensource.org/licenses/Apache-2.0>.
13. PyPI · the Python Package Index. *PyPI* <https://pypi.org/>.
14. Conda — Conda documentation. <https://docs.conda.io/en/latest/>.
15. Hornik, K. The comprehensive R archive network. *Wiley Interdiscip. Rev. Comput. Stat.* 4, 394–398 (2012).
16. Dockstore. <https://dockstore.org/>.
17. ModelDB: Home. <https://senselab.med.yale.edu/ModelDB/>.
18. McDougal, R. A. *et al.* Twenty years of ModelDB and beyond: building essential modeling tools for the future of neuroscience. *J. Comput. Neurosci.* 42, 1–10 (2017).

19. Bioconductor - home. <https://www.bioconductor.org/>.
20. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12, 115–121 (2015).
21. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5, R80 (2004).
22. Registry of Research Data Repositories. <https://www.re3data.org/>.